

## PATENT ABSTRACTS OF JAPAN

(11)Publication number: 09325798

(43)Date of publication of application: 16.12.1997

(51)Int.Cl.

G10L	9/00
G10L	3/02
G10L	3/02

(21)Application number: 08144728

(71)Applicant:

MATSUSHITA ELECTRIC IND CO LTD

(22)Date of filing: 06.06.1996

(72)Inventor:

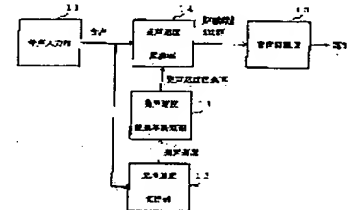
TERADA YASUHIRO

(54) VOICE RECOGNIZING DEVICE

(57)Abstract:

**PROBLEM TO BE SOLVED:** To improve a recognizing rate by converting (adjusting) an input voice in accordance with a voice uttering speed and absorbing variation of a voice uttering speed without changing a recognition processing section.

SOLUTION: An uttering speed of an input voice is calculated by an uttering speed calculating section 12 for a voice inputted from a voice input section 11, a conversion rate of an uttering speed for an input voice is decided by an uttering speed conversion rate deciding section 13 based on the calculated uttering speed, an uttering speed of an input voice is converted in a waveform level, the converted input voice is inputted to a voice recognizing section 15, and the recognized result is obtained.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-325798

(43) 公開日 平成9年(1997)12月16日

(51) Int.Cl. <sup>6</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 9/00	3 0 1		G 1 0 L 9/00	3 0 1 A
				3 0 1 B
3/02			3/02	A
	3 0 1			3 0 1 A

審査請求 未請求 請求項の数 4 O L (全 5 頁)

(21) 出願番号 特願平8-144728

(22) 出願日 平成8年(1996)6月6日

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 寺 田 泰 宏

神奈川県横浜市港北区綱島東四丁目3番1号 松下通信工業株式会社内

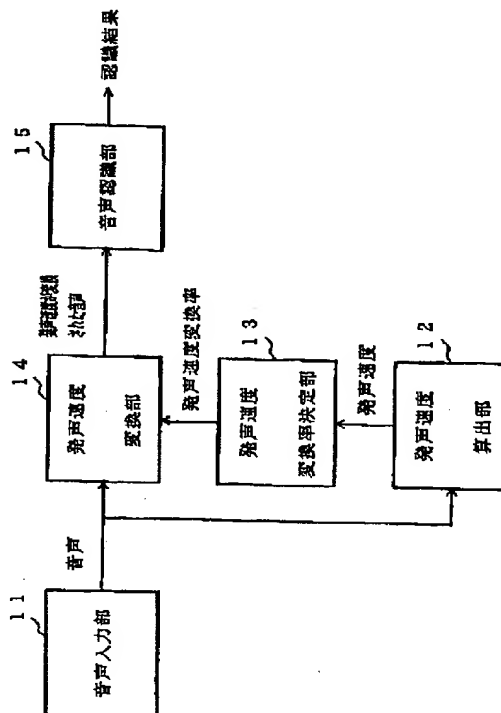
(74) 代理人 弁理士 蔵合 正博

(54) 【発明の名称】 音声認識装置

(57) 【要約】

【課題】 入力音声を発声速度に応じて変換（調整）することによって、認識処理部を変更せずに発声速度の変動を吸収し認識率の向上を図る。

【解決手段】 音声入力部11から入力された音声に対し、発声速度算出部12で入力音声の発声速度を算出し、その算出された発声速度をもとに発声速度変換率決定部13において入力音声に対する発声速度の変換率を決定し、その変換率をもとに発声速度変換部14において、入力音声の発声速度を波形レベルで変換し、変換された入力音声の音声認識部15に入力し、認識結果を得る。



## 【特許請求の範囲】

【請求項1】 音声を入力する音声入力部と、音声のパワーやピッチ周期等から入力音声の発声速度を算出する発声速度算出部と、算出された入力音声の発声速度を目標とする発声速度に変換する割合を決定する発声速度変換率決定部と、決定された変換率をもとに入力音声の発声速度に応じて時間長伸縮処理を行なって入力音声の発声速度を波形レベルで変換する発声速度変換部と、発声速度変換された音声を認識処理する音声認識部とを備えた音声認識装置。

【請求項2】 発声速度変換率決定部が、入力音声の発声速度を標準パターンを作成する際に利用した音声の発声速度の平均値になるように変換率を決定することを特徴とする請求項1記載の音声認識装置。

【請求項3】 発声速度変換率決定部が、入力音声の発声速度を音声認識処理部が対応する発声速度の許容範囲に入るように変換率を決定することを特徴とする請求項1記載の音声認識装置。

【請求項4】 発声速度変換における時間長伸縮処理において、入力音声の発声速度の調整にポイント移動量制御による重複加算法を用いることを特徴とする請求項1から3のいずれかに記載の音声認識装置。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、音声認識に関するものであり、発声速度の影響を軽減し、認識率の向上を図る音声認識装置に関する。

## 【0002】

【従来の技術】従来、音声認識における発声速度の影響を軽減する方法として、HMM法に基づいた音声認識を行なう場合には、入力音声の発声速度を検出し、標準パターンとのマッチングを行う際に、検出された発声速度を用いてHMMのある状態における分析フレームの継続時間長を調整しながらマッチングを行う方法が一般的である。一例を挙げると、特開平2-113298号公報に開示されているHMM音声認識におけるHMMの各状態の継続時間の制限を行う方法がそれであり、その構成を図5に示す。図5において、51は発声速度検出部、52は音声記号化部、53は単語HMMデータベース、54はHMM法に基づく音声認識部である。具体的な処理は次の通りである。入力端子から入力された音声は、発声速度検出部51で音声パワー、スペクトルから発声速度が求められ、HMM法に基づく音声認識部54に与えられる。音声記号化部52では、入力音声を記号系列に変換し、HMM法に基づく音声認識部54に与える。HMM法に基づく音声認識部54では、発声速度の値を用いて継続時間の制限を行いながら、記号系列と単語HMMデータベース53中の各単語HMMとの間の確率計算を行い、最も高い確率が得られるHMMに対応する単語を認識結果として出力する。

## 【0003】

【発明が解決しようとする課題】しかしながら、上記従来の技術では、認識処理を行う時点で発声速度の吸収を行うため、既存の音声認識装置または音響分析を含めた音声認識の方法や標準パターン等を改良しなければならないという課題があった。

【0004】本発明は、上記従来技術の課題を解決するものであり、音声認識装置または音響分析を含めた音声認識の方法や標準パターン等を改良することなく、発声速度の影響を軽減することのできる音声認識装置を提供することを目的としている。

## 【0005】

【課題を解決するための手段】本発明は、上記目的を達成するために、音声入力部、発声速度算出部、発声速度変換率決定部、発声速度変換部、音声認識部を備え、音声入力部から入力された音声に対して、発声速度算出部で入力音声の発声速度を算出し、算出した発声速度をもとに発声速度変換率決定部で入力音声の所望の発声速度になるような変換率を決定し、さらに発声速度変換部において波形レベルで入力音声の発声速度を変換し、発声速度を変換した音声を音声認識部に入力し、音声認識部において音声認識を行うようにしたものである。

## 【0006】

【発明の実施の形態】本発明の請求項1に記載の発明は、音声を入力する音声入力部と、音声のパワーやピッチ周期等から入力音声の発声速度を算出する発声速度算出部と、算出された入力音声の発声速度を目標とする発声速度に変換する割合を決定する発声速度変換率決定部と、決定された変換率をもとに入力音声の発声速度に応じて時間長伸縮処理を行なって入力音声の発声速度を波形レベルで変換する発声速度変換部と、発声速度変換された音声を認識処理する音声認識部とを備えた音声認識装置であり、入力音声自体の発声速度を変換するので、既存の音声認識装置または音声認識処理の方法や標準パターン等を改良せずに、発声速度の影響を軽減し認識性能が向上するという作用を有する。

【0007】本発明の請求項2に記載の発明は、発声速度変換率決定部が、入力音声の発声速度を標準パターンを作成する際に利用した音声の発声速度の平均値になるように変換率を決定することを特徴とする請求項1記載の音声認識装置であり、入力音声の発声速度を認識装置にとって最も認識しやすい発声速度に変換するので、認識性能が向上するという作用を有する。

【0008】本発明の請求項3に記載の発明は、発声速度変換率決定部が、入力音声の発声速度を音声認識処理部が対応する発声速度の許容範囲に入るように変換率を決定することを特徴とする請求項1記載の音声認識装置であり、入力音声の発声速度の変換に幅を持たせることにより、極端な発声速度の変換を防ぐことができ、入力音声の劣化を最小に発声速度の変換ができるので、発声

速度の変換が認識性能に影響を及ぼさないという作用を有する。

【0009】本発明の請求項4に記載の発明は、発声速度変換部における時間長伸縮処理において、入力音声の発声速度の調整にポイント移動量制御による重複加算法を用いることを特徴とする請求項1から3のいずれかに記載の音声認識装置であり、入力音声の自然性を保ったまま、入力音声の発声速度の変換が可能なので、認識性能に影響を及ぼさずに発声速度の変換が可能であるという作用を有する。

【0010】(実施の形態)以下、本発明の実施の形態について、図1～図4を用いて説明する。図1は本発明の実施の形態における音声認識装置の構成を示すものであり、図2は図1における発声速度算出部を実現する構成図の一例であり、図3は図1における発声速度変換部を実現する発声速度の変換方法であるポイント移動量制御による重複加算法の原理の説明図であり、図4は図1における音声認識部を実現する構成図の一例である。図1において、11は音声入力部、12は発声速度算出部、13は発声速度変換率決定部、14は発声速度変換部、15は音声認識部である。

【0011】以下、本実施の形態における動作について説明する。図1において、音声入力部11から入力された音声に対して、発声速度算出部12で、音声のパワーやピッチ周期等から入力音声の発声速度を算出する。次に、求めた発声速度をもとに発声速度変換率決定部13において入力音声に対し発声速度を変換する割合を決定する。例えば、発声速度算出部12で算出された発声速度が6.8音節/秒であり、目標とする発声速度が4.6音節/秒であったとしたら、発声速度変換率は148%(=6.8/4.6\*100)となる。この変換率を用いて発声速度変換部14において、入力音声の発声速度を波形レベルで変換、すなわち入力音声に対し148%の時間軸での伸張を行う。そして、発声速度が変換された音声は、音声認識部15に入力され認識処理が行われる。

【0012】図2は発声速度算出部12の構成を示している。図2において、入力音声は音声パワー算出部21に与えられ、一定長のフレーム毎のパワーが算出される。また、入力音声はピッチ周期算出部22にも与えられ、自己相関法、変形自己相関法等を用いてフレーム毎のピッチ周期が算出される。そして、母音/子音判定部23において、音声パワー算出部21から算出されたフレーム毎の音声パワーと、ピッチ周期算出部22から算出されたフレーム毎のピッチ周期とから、入力音声に対し母音/子音判定を行い、入力音声の中の母音の数から発声速度を求める。

【0013】なお、発声速度の算出は、予め決められた単語を発声してもらい、単語の音節数と発声音声の時間長から計算してもよい。

【0014】発声速度変換率決定部13においては、入

力音声に対する発声速度を標準パターンを作成する際に用いた音声の発声速度の平均値に調整するように変換率を決定する。具体的には、入力音声の発声速度が6.8音節/秒であり、標準パターンを作成する際に用いた音声の発声速度の平均値が4.4音節/秒であったとき、発声速度の変換率を130%(=6.8/5.2\*100)に設定し、図1に示した発声速度変換部14で、この値をもとに発声速度を変換する。

【0015】発声速度変換率決定部13において、変換率を決定する別の方法として、入力音声に対する発声速度を音声認識部15が対応する発声速度の許容範囲、具体的に例を挙げると、標準パターンを作成する際に用いた音声の発声速度の平均値を中心として標準偏差の範囲に入力音声の発声速度が変換されるように変換率を決定することができる。例えば、標準パターンを作成する際に用いた音声の発声速度の平均値が5.2音節/秒、標準偏差が1.1とした場合、音声認識部15が対応する発声速度の許容範囲は、4.1～6.3音節/秒となる。入力音声の発声速度が6.8音節/秒だとすると、発声速度変換率決定部13では、入力音声の劣化ができるだけ少なくなるような発声速度6.3音節/秒になるように変換率を決定する。具体的な数値は、108%(=6.8/6.3\*100)となる。この値をもとに発声速度変換部14において入力音声の発声速度を変換する。

【0016】発声速度変換部14における処理は、図3に示すようにして行なわれる。具体的には、入力音声に対し、ポイント移動量制御による重複加算法を用いて時間軸で圧縮・伸張を行う。ここでいう圧縮とは発声速度を早くすることであり、伸張とは発声速度を遅くすることである。まず、図3(a)を用いて圧縮の原理について説明する。現在処理を行うポイント(▽)から1ピッチ周期分のピッチ波形に対し、1から0へ直線的に向かう重み付けをする。そして、その隣り合うピッチ波形に対して、0から1へ直線的に向かう重み付けをし、両者を重ね合わせる。次に、2期分のピッチ波形を重ね合わせた波形に置き換える。そして、圧縮率をR(R<1)としたとき、現在処理を行うポイント(▽)からポイントの移動量 $L_c = R T_p / (1 - R)$ に従って次に処理を行うポイント(▼)を決定する。図3(a)はR=0.75のときの処理を示す。次に、図3(b)を用いて伸長の原理について説明する。現在処理を行うポイント(▽)から1ピッチ周期分のピッチ波形に対し、0から1へ直線的に向かう重み付けをする。そして、その隣り合うピッチ波形に対して、1から0へ直線的に向かう重み付けをし、両者を重ね合わせる。次に、重ね合わせた波形を2つのピッチ波形の間に挿入する。そして、伸張率をR(R>1)としたとき、挿入した波形の先頭のパポイント(▽)からポイントの移動量 $L_s = T_p / (R - 1)$ に従って次に処理を行うポイント(▼)を決定する。図3(b)はR=1.25のときの処理を示す。ここで、

$T_p$  はピッチ周期を示す。

【0017】このように、ポインタ移動量制御による重複加算法は処理を隣り合った2周期分のピッチ波形に限定するため、発声速度変換後の音質の劣化が少ないため、認識性能に影響を与えずに入力音声の発声速度の変換が可能となる。

【0018】図4は音声認識部15の構成を示している。図4において、入力音声に対し、フレーム毎に音響分析部41において音響分析を行い、入力音声をLPC ケプストラム、LPC メルケプストラム等のパラメータに変換する。変換された音声は、パターン照合部42においてDP (Dynamic Programing, 動的計画) マッチング法やHMM (Hidden Markov Model, 隠れマルコフモデル) 法などを用いて、フレーム毎に単語標準パターン格納部43に格納されている単語標準パターンと照合を行い、入力音声と単語標準パターンが最も近いものを認識結果として出力する。

【0019】

【発明の効果】以上のように、本発明によれば、音声入力部、発声速度算出部、発声速度変換率決定部、発声速度変換部、音声認識部を備え、音声入力部から入力された音声に対して、発声速度算出部で入力音声の発声速度を算出し、算出した発声速度をもとに発声速度変換率決定部で入力音声の所望の発声速度になるような変換率を決定し、さらに発声速度変換部において波形レベルで入力音声の発声速度を変換し、発声速度を変換した音声を音声認識部に入力し、音声認識部において音声認識を行うようにしたので、音声認識装置または音響分析を含めた音声認識の方法や標準パターン等を改良せずに発声速度の影響を軽減することが可能となる。

【図面の簡単な説明】

【図1】本発明の実施の形態における音声認識装置の構成を示すブロック図

【図2】本発明の実施の形態における発声速度算出部の構成を示すブロック図

【図3】(a) 本発明の実施の形態における発声速度変換部の発声速度を速くする時の原理を説明する波形図

(b) 本発明の実施の形態における発声速度変換部の発声速度を遅くする時の原理を説明する波形図

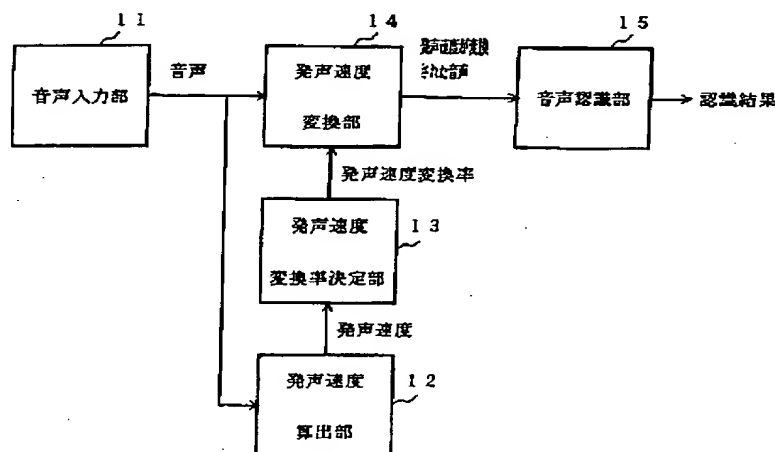
【図4】本発明の実施の形態における音声認識部の構成を示すブロック図

【図5】従来例における音声認識装置の構成を示すブロック図

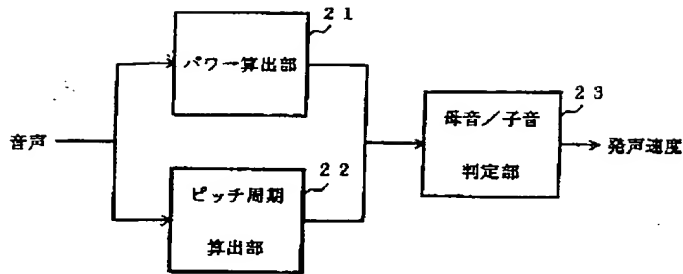
【符号の説明】

- 11... 音声入力部
- 12... 発声速度算出部
- 13... 発声速度変換率決定部
- 14... 発声速度変換部
- 15... 音声認識部
- 21... パワー算出部
- 22... ピッチ周期算出部
- 23... 母音／子音判定部
- 41... 音響分析部
- 42... パターン照合部
- 43... 単語標準パターン格納部
- 50... 入力端子
- 51... 発声速度検出部
- 52... 音声記号化部
- 53... 単語HMMデータベース
- 54... HMM法に基づく音声認識部

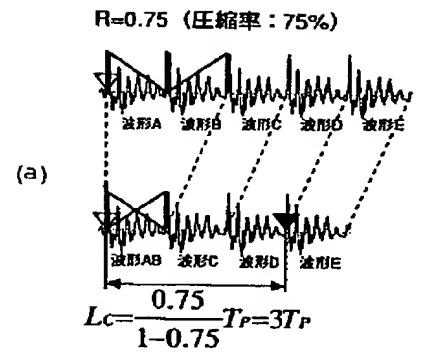
【図1】



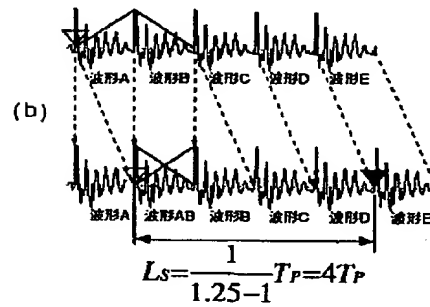
【図 2】



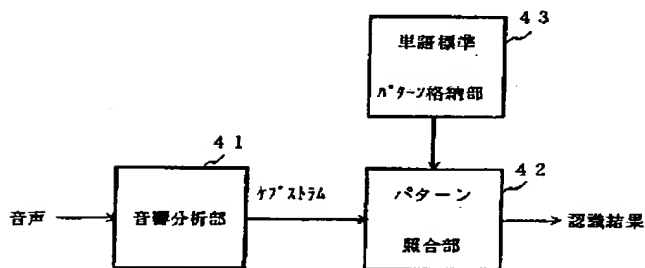
【図 3】



R=1.25 (伸長率: 125%)



【図 4】



【図 5】

